Copositive Programming – a Survey

Mirjam Dür

Johann Bernoulli Institute of Mathematics and Computer Science, University of Groningen, P.O. Box 407, 9700 AK Groningen, The Netherlands. M.E.Dur@rug.nl

Summary. Copositive programming is a relatively young field in mathematical optimization. It can be seen as a generalization of semidefinite programming, since it means optimizing over the cone of so called copositive matrices. Like semidefinite programming, it has proved particularly useful in combinatorial and quadratic optimization. The purpose of this survey is to introduce the field to interested readers in the optimization community who wish to get an understanding of the basic concepts and recent developments in copositive programming, including modeling issues and applications, the connection to semidefinite programming and sum-of-squares approaches, as well as algorithmic solution approaches for copositive programs.

1 Introduction

A copositive program is a linear optimization problem in matrix variables of the following form:

$$\begin{aligned} \min\ & \langle C, X\rangle \\ \text{s. t. } & \langle A_i, X\rangle = b_i \quad (i = 1, \dots, m), \\ & X \in \mathcal{C}, \end{aligned} \tag{1}$$

where $\mathcal{C}$ is the cone of so-called copositive matrices, that is, the matrices whose quadratic form takes nonnegative values on the nonnegative orthant $\mathbb{R}^n_+$:

$$\mathcal{C} = \{A \in \mathcal{S} : x^T Ax \geq 0 \text{ for all } x \in \mathbb{R}^n_+\}$$

(here $\mathcal{S}$ is the set of symmetric $n \times n$ matrices, and the inner product of two matrices in (1) is $\langle A, B\rangle := \text{trace}(BA) = \sum_{i,j=1}^n a_{ij} b_{ij}$, as usual). Obviously, every positive semidefinite matrix is copositive, and so is every entrywise nonnegative matrix, but the copositive cone is significantly larger than both the semidefinite and the nonnegative matrix cones.

Interpreting (1) as the primal program, one can associate a corresponding dual program which is a maximization problem over the dual cone. For an arbitrary given cone $\mathcal{K} \subseteq \mathcal{S}$, the dual cone $\mathcal{K}^*$ is defined as

M. Diehl et al. (eds.), *Recent Advances in Optimization and its Applications in Engineering*, DOI 10.1007/978-3-642-12598-0_1, © Springer-Verlag Berlin Heidelberg 2010

$$\mathcal{K}^* := \{A \in \mathcal{S} : \langle A, B \rangle \geq 0 \text{ for all } B \in \mathcal{K}\}.$$

In contrast to the semidefinite and nonnegative matrix cones, the cone $\mathcal{C}$ is not selfdual. It can be shown (see e.g. [6]) that $\mathcal{C}^*$ is the cone of so-called completely positive matrices

$$\mathcal{C}^* = \operatorname{conv}\{xx^T : x \in \mathbb{R}^n_+\}.$$

Using this, the dual of (1) can be derived through the usual Lagrangian approach and is easily seen to be

$$\begin{aligned} \max & \sum_{i=1}^m b_i y_i \\ \text{s.t. } & C - \sum_{i=1}^m y_i A_i \in \mathcal{C}^*, \; y_i \in \mathbb{R}. \end{aligned} \tag{2}$$

Since both $\mathcal{C}$ and $\mathcal{C}^*$ are convex cones, (1) and (2) are convex optimization problems. KKT optimality conditions hold if Slater's condition is satisfied, as shown by [28], and imposing a constraint qualification guarantees strong duality, i.e., equality of the optimal values of (1) and (2). The most common constraint qualification is to assume that both problems are feasible and one of them strictly feasible (meaning that there exists a strictly feasible point, i.e., a solution to the linear constraints in the interior of the cone).

Copositive programming is closely related to quadratic and combinatorial optimization. We illustrate this connection by means of the standard quadratic problem

$$(\text{StQP}) \qquad \begin{aligned} \min\ & x^T Q x \\ \text{s.t. } & e^T x = 1, \\ & x \geq 0, \end{aligned}$$

where e denotes the all-ones vector. This optimization problem asks for the minimum of a (not necessarily convex) quadratic function over the standard simplex. Easy manipulations show that the objective function can be written as $x^T Q x = \langle Q, xx^T \rangle$. Analogously the constraint $e^T x = 1$ transforms to $\langle E, xx^T \rangle = 1$, with $E = ee^T$. Hence, the problem

$$\begin{aligned} \min\ & \langle Q, X \rangle \\ \text{s.t. } & \langle E, X \rangle = 1, \\ & X \in \mathcal{C}^* \end{aligned} \tag{3}$$

is obviously a relaxation of (StQP). Since the objective is now linear, an optimal solution must be attained in an extremal point of the convex feasible set. It can be shown that these extremal points are exactly the rank-one matrices xx^T with $x \geq 0$ and $e^T x = 1$. Together, these results imply that (3) is in fact an exact reformulation of (StQP).

The standard quadratic problem is an NP-hard optimization problem, since the maximum clique problem can be reduced to an (StQP). Indeed, denoting by $\omega(G)$ the clique number of a graph G and by A_G its adjacency matrix, Motzkin and Straus [43] showed that

$$\frac{1}{\omega(G)} = \min\{x^T(E - A_G)x : e^T x = 1, x \geq 0\}. \tag{4}$$

Nevertheless, (3) is a convex formulation of this NP-hard problem. This shows that NP-hard convex optimization problems do exist. The complexity has moved entirely into the cone-constraint $X \in \mathcal{C}^*$. It is known that testing whether a given matrix is in $\mathcal{C}$ is co-NP-complete (cf. [44]). Consequently, it is not tractable to do a line-search in $\mathcal{C}$. The cones $\mathcal{C}$ and $\mathcal{C}^*$ do allow self-concordant barrier functions (see [46]), but these functions can not be evaluated in polynomial time. Thus, the classical interior point methodology does not work. Optimizing over either $\mathcal{C}$ or $\mathcal{C}^*$ is thus NP-hard, and restating a problem as an optimization problem over one of these cones does not resolve the difficulty of that problem. However, studying properties of $\mathcal{C}$ and $\mathcal{C}^*$ and using the conic formulations of quadratic and combinatorial problems does provide new insights and also computational improvements.

Historical remarks

The concept of copositivity seems to go back to Motzkin [42] in the year 1952. Since then, numerous papers on both copositivity and complete positivity have emerged in the linear algebra literature, see [6] or [36] for surveys. Using these cones in optimization has been studied only in the last decade.

An early paper relating the solution of a certain quadratic optimization problem to copositivity is Preisig [52] from 1996. Preisig describes properties and derives an algorithm for what we would now call the dual problem of (3) with E replaced by a strictly copositive matrix B. However, he just analyzes this particular problem and does not provide the conic programming framework outlined above. It seems that his paper has been widely ignored by the optimization community.

Quist et al. [53] suggested in 1998 that semidefinite relaxations of quadratic problems may be tightened by looking at the copositive cone. They were the first to formulate problems with the conic constraints $X \in \mathcal{C}$ and $X \in \mathcal{C}^*$.

Bomze et al. [11] were the first to establish an equivalent copositive formulation of an NP-hard problem, namely the standard quadratic problem. Their paper from 2000 also coined the term "copositive programming".

Since [11] appeared, a number of other quadratic and combinatorial problems have been shown to admit an exact copositive reformulation. Although these formulations remain NP-hard, they have inspired better bounds than previously available. Through sum-of-squares approximations (cf. Section 5 below) they have opened a new way to solve these problems. Finally, new solution algorithms for copositive and completely positive problems have been developed and proved very successful in some settings, as we describe in Section 6.

2 Applications

Binary quadratic problems

We have seen in Section 1 that the standard quadratic problem can be rewritten as a completely positive program. This can be extended to so-called multi-StQPs, where one seeks to optimize a quadratic form over the cartesian product of simplices, see [15].

Burer [19] showed the much more general result that every quadratic problem with linear and binary constraints can be rewritten as such a problem. More precisely, he showed that a quadratic binary problem of the form

$$\begin{aligned} \min\ & x^TQx + 2c^Tx \\ \text{s.t. } & a_i^Tx = b_i \quad (i = 1, \dots, m) \\ & x \geq 0 \\ & x_j \in \{0,1\} \quad (j \in B) \end{aligned} \tag{5}$$

can equivalently be reformulated as the following completely positive problem:

$$\begin{aligned} \min\ & \langle Q, X\rangle + 2c^Tx \\ \text{s.t. } & a_i^Tx = b_i \quad (i = 1, \dots, m) \\ & \langle a_ia_i^T, X\rangle = b_i^2 \quad (i = 1, \dots, m) \\ & x_j = X_{jj} \quad (j \in B) \\ & \begin{pmatrix} 1 & x \\ x & X \end{pmatrix} \in \mathcal{C}^*, \end{aligned}$$

provided that (5) satisfies the so-called key condition, i.e., $a_i^Tx = b_i$ for all i and $x \geq 0$ implies $x_j \leq 1$ for all $j \in B$. As noted by Burer, this condition can be enforced without loss of generality.

It is still an open question whether problems with general quadratic constraints can similarly be restated as completely positive problems. Only special cases like complementarity constraints have been solved [19]. For a comment on Burer's result see [13]. Natarajan et al. [45] consider (5) in the setting where $Q = 0$ and c is a random vector, and derive a completely positive formulation for the expected optimal value.

Fractional quadratic problems

Consider a matrix A whose quadratic form x^TAx does not have zeros in the standard simplex, i.e., consider without loss of generality a strictly copositive matrix A. Preisig [52] observed that then the problem of maximizing the ratio of two quadratic forms over the standard simplex

$$\min\left\{\frac{x^TQx}{x^TAx} : e^Tx = 1, x \geq 0\right\}$$

is equivalent to

$$\min\{x^T Qx : x^T Ax = 1, x \geq 0\}$$

and hence, by similar arguments as used to derive (3), is equivalent to the completely positive program

$$\min\{\langle Q, X\rangle : \langle A, X\rangle = 1, x \in \mathcal{C}^*\}$$

For a thorough discussion, see also [16].

Combinatorial problems

For the problem of determining the **clique number** $\omega(G)$ of a graph G, we can combine the Motzkin-Straus formulation (4) with the completely positive formulation (3) of the standard quadratic problem. Taking the dual of that problem, we arrive at

$$\tfrac{1}{\omega(G)} = \max\{\lambda : \lambda(E - A_G) - E \in \mathcal{C}\}.$$

Using a somewhat different approach, De Klerk and Pasechnik [23] derive the following formulation for the **stability number** $\alpha(G)$:

$$\alpha(G) = \min\{\lambda : \lambda(I + A_G) - E \in \mathcal{C}\}$$

(I the identity matrix), or, in the dual formulation,

$$\alpha(G) = \max\{\langle E, X\rangle : \langle A_G + I, X\rangle = 1, X \in \mathcal{C}^*\}.$$

The last formulation can be seen as a strengthening of the Lovász ϑ number, which is obtained by optimizing over the cone $\mathcal{S}^+ \cap \mathcal{N}$ of entrywise nonnegative and positive semidefinite matrices instead of $\mathcal{C}^*$ in the above problem.

Dukanovic and Rendl [26] introduce a related copositivity-inspired strengthening of the Lovász ϑ number toward the chromatic number of G, which is shown to be equal to the fractional chromatic number.

For the **chromatic number** $\chi(G)$ of a graph G with n nodes, a copositive formulation has been found by Gvozdenović and Laurent in [30]:

$$\begin{aligned} \chi(G) = \max\ & y \\ \text{s. t. }\ & \tfrac{1}{n^2}(ty)E + z(n(I + A_{G_t})E) \in \mathcal{C} \qquad t = 1, \dots, n \\ & y, z \in \mathbb{R}. \end{aligned}$$

where A_{G_t} denotes the adjacency matrix of the graph G_t, the cartesian product of the graphs K_t (the complete graph on t nodes) and G. This product graph G_t has node set $V(K_t) \times V(G) = \bigcup_{p=1}^{t} V_p$, where $V_p := \{pi : i \in V(G)\}$. An edge (pi, qj) is present in G_t if ($p \neq q$ and $i = j$) or if ($p = q$ and (ij) is an edge in G).

A completely positive formulation of the related problem of computing the **fractional chromatic number** can be found in [26].

A completely positive formulation for the **quadratic assignment problem** (QAP) was developed in [50]. Introducing it requires some notation: let A, B, C be the matrices describing the QAP instance. $B \otimes A$ denotes the Kronecker product of B and A, i.e., the $n^2 \times n^2$ matrix $(b_{ij}A)$. Let $c = \mathrm{vec}(C)$ be the vector derived from C by stacking the columns of C on top of each other, and let $\mathrm{Diag}(c)$ be the $n^2 \times n^2$ diagonal matrix with the entries of c on its diagonal. The variable Y of the completely positive problem is also an $n^2 \times n^2$ matrix. Its $n \times n$ component blocks are addressed by Y^{ij} with $i, j = 1, \ldots, n$. Finally, δ_{ij} is the Kronecker-delta.

Using this notation, Povh and Rendl [50] show that the optimal value of QAP is the solution of the following completely positive program of order n^2:

$$\begin{aligned} OPT_{QAP} = \min\ & \langle B \otimes A + \mathrm{Diag}(c), Y \rangle \\ \text{s.t. } & \textstyle\sum_i Y^{ii} = I \\ & \langle I, Y^{ij} \rangle = \delta_{ij} \quad (i, j = 1, \ldots, n) \\ & \langle E, Y \rangle = n^2 \\ & Y \in \mathcal{C}^*. \end{aligned}$$

The problem of finding a **3-partitioning** of the vertices of a graph G was studied by Povh and Rendl in [51]. Consider a graph on n vertices with weights $a_{ij} \geq 0$ on its edges. The problem is to partition the vertices of G into subsets S_1, S_2, and S_3 with given cardinalities m_1, m_2, and m_3 (with $\sum_i m_i = n$) in such a way that the total weight of edges between S_1 and S_2 is minimal. Note that this problem contains the classical graph bisection problem as a special case.

The completely positive formulation requires some notation again. Letting e_i denote the ith unit vector in appropriate dimension, take $E_{ij} = e_i e_j^T$ and B_{ij} its symmetrized version $B_{ij} = 1/2(E_{ij} + E_{ji})$. For $j = 1, \ldots, n$, define matrices $W_j \in \mathbb{R}^{n \times n}$ by $W_j = e_j e^T$. Moreover, define the following 3×3 matrices: E_3 the all-ones matrix in $\mathbb{R}^{3\times3}$, $B = 2B_{12}$ in $\mathbb{R}^{3\times3}$ and for $i = 1, 2, 3$ define $V_i \in \mathbb{R}^{3\times3}$ as $V_i = e_i e^T$.

With these notations, Povh and Rendl derive the following completely positive formulation of order $3n$:

$$\begin{aligned} \min\ & \tfrac{1}{2}\langle B^T \otimes A, Y \rangle \\ \text{s.t. } & \langle B_{ij} \otimes I, Y \rangle = m_i \delta_{ij} \qquad 1 \leq i \leq j \leq 3 \\ & \langle E_3 \otimes E_{ii}, Y \rangle = 1 \qquad i = 1, \ldots, n \\ & \langle V_i \otimes W_j^T, Y \rangle = m \qquad i = 1, 2, 3;\ j = 1, \ldots, n \\ & \langle B_{ij} \otimes E, Y \rangle = m_i m_j \qquad 1 \leq i \leq j \leq 3 \\ & Y \in \mathcal{C}^*. \end{aligned}$$

As far as we are aware, the above list comprises all problem classes for which an equivalent copositive or completely positive formulation has been established up to now. It illustrates that copositive programming is a powerful modelling tool which interlinks the quadratic and binary worlds. In the next sections, we will discuss properties of the cones as well as algorithmic approaches to tackle copositive programs.

3 The cones $\mathcal{C}$ and $\mathcal{C}^*$

Topological properties

Both $\mathcal{C}$ and $\mathcal{C}^*$ are full-dimensional closed, convex, pointed, non-polyhedral matrix cones. The interior of $\mathcal{C}$ is the set of strictly copositive matrices:

$$\operatorname{int}(\mathcal{C}) = \{A : x^T Ax > 0 \text{ for all } x \geq 0, x \neq 0\}.$$

The extremal rays of $\mathcal{C}^*$ are the rank-one completely positive matrices

$$\operatorname{Ext}(\mathcal{C}^*) = \{xx^T : x \geq 0\}.$$

Proofs of all these statements can be found in [6]. The interior of the completely positive cone has first been characterized in [27]. Dickinson [24] gave an improved characterization which reads as follows:

$$\operatorname{int}(\mathcal{C}^*) = \{AA^T : \operatorname{rank}(A) = n \text{ and } A = [a|B] \text{ with } a \in \mathbb{R}^n_{++}, B \geq 0\}.$$

Here the notation $[a|B]$ describes the matrix whose first column is a and whose other columns are the columns of B. An alternative characterization is

$$\operatorname{int}(\mathcal{C}^*) = \{AA^T : \operatorname{rank}(A) = n \text{ and } A > 0\}.$$

A full characterization of the extremal rays of $\mathcal{C}$ (or equivalently, a complete "outer" description of $\mathcal{C}^*$ in terms of supporting hyperplanes) is an open problem. Partial results can be found in [3, 4, 5, 32, 34].

Small dimensions

The cones $\mathcal{C}$ and $\mathcal{C}^*$ are closely related to the cones $\mathcal{S}^+$ of positive semidefinite matrices and $\mathcal{N}$ of entrywise nonnegative matrices, since we immediately get from the definitions that

$$\mathcal{C}^* \subseteq \mathcal{S}^+ \cap \mathcal{N} \quad \text{and} \quad \mathcal{C} \supseteq \mathcal{S}^+ + \mathcal{N}.$$

Matrices in $\mathcal{S}^+ \cap \mathcal{N}$ are sometimes called doubly nonnegative. It is a very interesting fact (cf. [41]) that for $n \times n$-matrices of order $n \leq 4$, we have equality in the above relations, whereas for $n \geq 5$, both inclusions are strict. A counterexample that illustrates $\mathcal{C} \neq \mathcal{S}^+ + \mathcal{N}$ is the so-called Horn-matrix, cf. [31]:

$$H = \begin{pmatrix} 1 & -1 & 1 & 1 & -1 \\ -1 & 1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & -1 & 1 \end{pmatrix}.$$

To see that H is copositive, write

$$\begin{aligned} x^T Hx &= (x_1 - x_2 + x_3 + x_4 - x_5)^2 + 4x_2x_4 + 4x_3(x_5 - x_4) \\ &= (x_1 - x_2 + x_3 - x_4 + x_5)^2 + 4x_2x_5 + 4x_1(x_4 - x_5). \end{aligned}$$

The first expression shows that $x^T Hx \geq 0$ for nonnegative x with $x_5 \geq x_4$, whereas the second shows $x^T Hx \geq 0$ for nonnegative x with $x_5 < x_4$. It can be shown [31] that H is extremal for $\mathcal{C}$, and consequently H can not be decomposed into $H = S + N$ with $S \in \mathcal{S}^+$ and $N \in \mathcal{N}$.

Why is this jump when the size of A changes from 4×4 to 5×5? This question was answered by Kogan and Berman [40] using graph theoretic arguments: associate to a given symmetric matrix $A \in \mathbb{R}^{n \times n}$ a graph G with n vertices, such that an edge (i,j) is present in G if and only if $A_{ij} \neq 0$. Kogan and Berman [40] define a graph G to be completely positive, if every matrix $A \in \mathcal{S}^+ \cap \mathcal{N}$ whose graph is G is completely positive, and they show that a graph is completely positive if and only if it does not contain a long odd cycle, i.e., a cycle of length greater than 4. Obviously, this can not happen in graphs on four vertices, which shows that for small dimensions $\mathcal{C}^* = \mathcal{S}^+ \cap \mathcal{N}$. Observe that the Horn-matrix is related to the 5-cycle via $H = E - 2A_5$, where A_5 the adjacency matrix of the 5-cycle.

The case of 5×5 copositive and completely positive matrices has therefore attracted special interest, and several papers have dealt with this setting, see [20] and references therein.

4 Testing copositivity and complete positivity

Complexity

It has been shown by Murty and Kabadi [44] that checking whether a given matrix $A \in \mathcal{C}$ is a co-NP-complete decision problem. Intuitively, checking $A \in \mathcal{C}^*$ should have the same computational complexity. It seems, however, that a formal proof of this statement has not yet been given.

This general complexity result does not exclude that for special matrix classes checking copositivity is cheaper. For example, for diagonal matrices one only needs to verify nonnegativity of the diagonal elements, evidently a linear-time task. This can be generalized: For tridiagonal matrices [10] and for acyclic matrices [35], testing copositivity is possible in linear time.

Complete positivity

There are several conditions, necessary and sufficient ones, for complete positivity of a matrix. Most of them use linear algebraic arguments or rely on properties of the graph associated to the matrix, and it seems unclear how they can be used for algorithmic methods to solve optimization problems over $\mathcal{C}^*$. For a comprehensible survey of these conditions, we refer to [6]. We just mention two sufficient conditions: a sufficient condition shown in [39] is that A is nonnegative and diagonally dominant. Another sufficient condition for $A \in \mathcal{S}^+ \cap \mathcal{N}$ to be in $\mathcal{C}^*$ is that A is tridiagonal or acyclic, as shown in [8].

Decomposing a given matrix $A \in \mathcal{C}^*$ into $A = \sum_{i=1}^{k} b_i b_i^T$ is also a nontrivial task. Since this is equivalent to finding a nonnegative matrix $B \in \mathbb{R}^{n \times k}$ (whose columns are b_i) with $A = BB^T$, this is sometimes called nonnegative factorization of A. A major line of research in the linear algebra literature is concerned with determining the minimal number k of factors necessary in such a decomposition. This quantity is called the cp-rank, and is conjectured [25] to be $\lfloor n^2/4 \rfloor$ if n is the order of the matrix. See [6] for more details on the cp-rank. Berman and Rothblum [7] proposed a non-polynomial algorithm to compute the cp-rank (and thus to determine whether a matrix is completely positive). Their method, however, does not provide a factorization. Jarre and Schmallowsky [37] also propose a procedure which for a given matrix A either determines a certificate proving $A \in \mathcal{C}^*$ or converges to a matrix $S \in \mathcal{C}^*$ which is in some sense "close" to A. Bomze [9] shows how a factorization of A can be used to construct a factorization of $\begin{pmatrix} 1 & b^T \\ b & A \end{pmatrix}$.

Copositivity criteria based on structural matrix properties

Obviously, copositivity of a matrix can not be checked through its eigenvalues. It can be checked by means of the so-called Pareto eigenvalues [33], but computing those is not doable in polynomial time. Spectral properties of copositive matrices provide some information and are discussed in [38].

For dimensions up to four, explicit descriptions are available [33]. For example, a symmetric 2×2 matrix A is copositive if and only if its entries fulfill $a_{11} \geq 0, a_{22} \geq 0$ and $a_{12} + \sqrt{a_{11}a_{22}} \geq 0$, see [1]. As this description indicates, the boundary of the cone $\mathcal{C}$ has both "flat parts" and "curved parts", so the cone is neither polyhedral nor strictly nonpolyhedral everywhere. This geometry and the facial structure of $\mathcal{C}$ is, however, not well-understood.

In all dimensions, copositive matrices necessarily have nonnegative diagonal elements: if $a_{ii} < 0$ for some i, then the corresponding coordinate vector e_i would provide $e_i^T A e_i = a_{ii} < 0$, thus contradicting copositivity of A.

A condition similar to the Schur-complement also holds for copositive matrices, as shown in [29]: Consider

$$A = \begin{pmatrix} a & b^T \\ b & C \end{pmatrix}$$

with $a \in \mathbb{R}$, $b \in \mathbb{R}^n$ and $C \in \mathbb{R}^{n \times n}$. Then A is copositive iff $a \geq 0$, C is copositive, and $y^T(aC - bb^T)y \geq 0$ for all $y \in \mathbb{R}^n_+$ such that $b^T y \leq 0$.

Numerous criteria for copositivity in terms of structural properties of the matrix have been given, many of them in terms of properties of principal submatrices. We name just one example stated in [21] but attributed to Motzkin: a symmetric matrix is strictly copositive iff each principal submatrix for which the cofactors of the last row are all positive has a positive determinant. Many conditions of the same flavor can be found in the literature. Again, it seems doubtful whether those conditions will prove useful for optimization purposes, so we refer to the surveys [33] and [36] for a more thorough treatment.

A recursive method to determine copositivity of a matrix has been proposed by Danninger [22].

An algorithmic approach

A conceptually different approach to copositivity testing which essentially uses global optimization techniques has been proposed in [18]. This approach relies on the observation that A is copositive iff the quadratic form $x^TAx \geq 0$ on the standard simplex. If $v_1, \dots, v_n$ denote the vertices of a simplex, we can write a point x in the simplex in barycentric coordinates as $x = \sum_{i=1}^n \lambda_i v_i$ with $\lambda_i \geq 0$ and $\sum_{i=1}^n \lambda_i = 1$. This gives

$$x^TAx = \sum_{i,j=1}^n v_i^TAv_j\lambda_i\lambda_j.$$

Hence, a necessary condition for x^TAx to be nonnegative on the simplex is that

$$v_i^TAv_j \geq 0 \text{ for all } i, j. \tag{6}$$

This condition can be refined by studying simplicial partitions of the standard simplex. As the partition gets finer, stronger and stronger necessary conditions are derived which, in the limit, capture all strictly copositive matrices. This approach gives very good numerical results for many matrices. It can be generalized in such a way that cones between $\mathcal{N}$ and $\mathcal{S}^+ + \mathcal{N}$ are used as certificates, see [54].

5 Approximation hierarchies

A matrix is copositive if its quadratic form is nonnegative for nonnegative arguments. Based on this definition, various approaches have used conditions which ensure positivity of polynomials.

For a given matrix $A \in \mathcal{S}$, consider the polynomial

$$P_A(x) := \sum_{i=1}^n \sum_{j=1}^n a_{ij}x_i^2x_j^2.$$

Clearly, $A \in \mathcal{C}$ if and only if $P_A(x) \geq 0$ for all $x \in \mathbb{R}^n$. A sufficient condition for this is that $P_A(x)$ has a representation as a sum of squares (sos) of polynomials. Parrilo [47] showed that $P_A(x)$ allows a sum of squares decomposition if and only if $A \in \mathcal{S}^+ + \mathcal{N}$, yielding again the relation $\mathcal{S}^+ + \mathcal{N} \subseteq \mathcal{C}$.

A theorem by Pólya [49] states that if $f(x_1,\dots,x_n)$ is a homogeneous polynomial which is positive on the standard simplex, then for sufficiently large $r \in \mathbb{N}$ the polynomial

$$f(x_1,\dots,x_n)\cdot\left(\sum_{i=1}^n x_i^2\right)^r$$

has positive coefficients. Inspired by this result, Parrilo [47] (cf. also [23] and [12]) defined the following hierarchy of cones for $r \in \mathbb{N}$:

$$\mathcal{K}^r := \left\{ A \in \mathcal{S} : P_A(x)\left(\sum_{i=1}^n x_i^2\right)^r \text{ has an sos decomposition } \right\}.$$

Parrilo showed $\mathcal{S}^+ + \mathcal{N} = \mathcal{K}^0 \subset \mathcal{K}^1 \subset \dots$, and $\mathrm{int}(\mathcal{C}) \subseteq \bigcup_{r\in\mathbb{N}} \mathcal{K}^r$, so the cones $\mathcal{K}^r$ approximate $\mathcal{C}$ from the interior. Since the sos condition can be written as a system of linear matrix inequalities (LMIs), optimizing over $\mathcal{K}^r$ amounts to solving a semidefinite program.

Exploiting a different sufficient condition for nonnegativity of a polynomial, De Klerk and Pasechnik [23], cf. also Bomze and De Klerk [12], define

$$\mathcal{C}^r := \left\{ A \in \mathcal{S} : P_A(x)\left(\sum_{i=1}^n x_i^2\right)^r \text{ has nonnegative coefficients } \right\}.$$

De Klerk and Pasechnik showed that $\mathcal{N} = \mathcal{C}^0 \subset \mathcal{C}^1 \subset \dots$, and $\mathrm{int}(\mathcal{C}) \subseteq \bigcup_{r\in\mathbb{N}} \mathcal{C}^r$. Each of these cones is polyhedral, so optimizing over one of them is solving an LP.

Refining these approaches, Peña et al. [48] derive yet another hierarchy of cones approximating $\mathcal{C}$. Adopting standard multiindex notation, where for a given multiindex $\beta \in \mathbb{N}^n$ we have $|\beta| := \beta_1 + \dots + \beta_n$ and $x^\beta := x_1^{\beta_1}\cdots x_n^{\beta_n}$, they define the following set of polynomials

$$\mathcal{E}^r := \left\{ \sum_{\beta\in\mathbb{N}^n, |\beta|=r} x^\beta x^T(S_\beta + N_\beta)x : S_\beta \in \mathcal{S}^+, N_\beta \in \mathcal{N} \right\}.$$

With this, they define the cones

$$\mathcal{Q}^r := \left\{ A \in \mathcal{S} : x^T A x \left(\sum_{i=1}^n x_i^2\right)^r \in \mathcal{E}^r \right\}.$$

They show that $\mathcal{C}^r \subseteq \mathcal{Q}^r \subseteq \mathcal{K}^r$ for all $r \in \mathbb{N}$, with $\mathcal{Q}^r = \mathcal{K}^r$ for $r = 0, 1$. Similar to $\mathcal{K}^r$, the condition $A \in \mathcal{Q}^r$ can be rewritten as a system of LMIs. Optimizing over $\mathcal{Q}^r$ is therefore again an SDP.

All these approximation hierarchies approximate $\mathcal{C}$ uniformly and thus do not take into account any information provided by the objective function of an optimization problem. Moreover, in all these approaches the system of LMIs (resp. linear inequalities) gets large quickly as r increases. Thus, dimension of the SDPs increases so quickly that current SDP-solvers can only solve problems over those cones for small values of r, i.e., $r \leq 3$ at most.

We are not aware of comparable approximation schemes that approximate the completely positive cone $\mathcal{C}^*$ from the interior.

6 Algorithms

The approximation hierarchies described in the last section can be used to approximate a copositive program, and in many settings this gives very good results and strong bounds. However, the size of the problems increases exponentially as one goes through the approximation levels, so only low-level approximations are tractable.

As far as we are aware, there are two approaches to solve copositive programs directly: one is a feasible descent method in the completely positive cone $\mathcal{C}^*$, the other one approximates the copositive cone $\mathcal{C}$ by a sequence of polyhedral inner and outer approximations. In the sequel we briefly describe both methods.

Optimizing over $\mathcal{C}^*$

A recent attempt to solve optimization problems over $\mathcal{C}^*$ is a feasible descent method by Bomze et al. [14], who approximate the steepest descent path from a feasible starting point in $\mathcal{C}^*$. They study the problem

$$\begin{aligned} \min\ & \langle C, X\rangle \\ \text{s.t. } & \langle A_i, X\rangle = b_i \quad (i = 1, \dots, m), \\ & X \in \mathcal{C}^*. \end{aligned} \tag{7}$$

The optimal solution is approximated by a sequence of feasible solutions, and in this sense the algorithm resembles an interior point method. Starting from an initial feasible solution X^0 of which a factorization $X^0 = (V^0)(V^0)^T$ is assumed to be available, the next iteration point is $X^{j+1} = X^j + \Delta X^j$, where ΔX^j is a solution of the following regularized version of (7):

$$\begin{aligned} \min\ & \varepsilon\langle C, \Delta X\rangle + (1-\varepsilon)\|\Delta X\|_j^2 \\ \text{s.t. } & \langle A_i, \Delta X\rangle = 0 \quad (i = 1, \dots, m), \\ & X^j + \Delta X \in \mathcal{C}^*. \end{aligned}$$

The norm $\|\cdot\|_j$ used in iteration j depends on the current iterate X^j. Setting $X^{j+1} = (V + \Delta V)(V + \Delta V)^T$, they show the regularized problem to be equivalent to

$$\begin{aligned} \min\ & \varepsilon\langle C, V(\Delta V)^T + (\Delta V)V^T + (\Delta V)(\Delta V)^T\rangle \\ & \qquad +(1-\varepsilon)\|V(\Delta V)^T + (\Delta V)V^T + (\Delta V)(\Delta V)^T\|_j^2 \\ \text{s.t. } & \langle A_i, V(\Delta V)^T + (\Delta V)V^T + (\Delta V)(\Delta V)^T\rangle = 0 \quad (i = 1,\dots,m), \\ & V + \Delta V \in \mathcal{N}. \end{aligned}$$

This problem now involves the tractable cone $\mathcal{N}$ instead of $\mathcal{C}^*$, but the objective is now a nonconvex quadratic function, and the equivalence statement only holds for the global optimum. Using linearization techniques and Tikhonov regularization for this last problem in V-space, the authors arrive at an implementable algorithm which shows promising numerical performance for the max-clique problem as well as box-constrained quadratic problems.

Convergence of this method is not guaranteed. Moreover, the algorithm requires knowledge of a feasible starting point together with its factorization. Finding a feasible point is in general as difficult as solving the original problem, and given the point, finding the factorization is highly nontrivial. In special settings, however, the factorized starting point comes for free.

Optimizing over $\mathcal{C}$

An algorithm for the copositive optimization problem (1) has been proposed in [17]. We also refer to [16] for a detailed elaboration. The method is based on the copositivity conditions developed in [18] which we briefly described in Section 4. Recall condition (6). Consider a simplicial partition $\mathcal{P}$ of the standard simplex Δ into smaller simplices, i.e., a family $\mathcal{P} = \{\Delta^1, \dots, \Delta^m\}$ of simplices satisfying $\Delta = \bigcup_{i=1}^m \Delta^i$ and $\operatorname{int}(\Delta^i) \cap \operatorname{int}(\Delta^j) = \emptyset$ for $i \neq j$. We denote the set of all vertices of simplices in $\mathcal{P}$ by

$$V_{\mathcal{P}} = \{v : v \text{ is a vertex of some simplex in } \mathcal{P}\},$$

and the set of all edges of simplices in $\mathcal{P}$ by

$$E_{\mathcal{P}} = \{(u, v) : u \neq v \text{ are vertices of the same simplex in } \mathcal{P}\}.$$

In this notation, the necessary copositivity condition from [18] reads: a matrix A is copositive if $v^T Av \geq 0$ for all $v \in V_{\mathcal{P}}$ and $u^T Av \geq 0$ for all $(u, v) \in E_{\mathcal{P}}$, cf. (6). This motivates to define the following set corresponding to a given partition $\mathcal{P}$:

$$\begin{aligned} \mathcal{I}_{\mathcal{P}} := \{A \in \mathcal{S} : \ & v^T Av \geq 0 \ \text{ for all } v \in V_{\mathcal{P}}, \\ & u^T Av \geq 0 \ \text{ for all } (u, v) \in E_{\mathcal{P}}\}. \end{aligned}$$

It is not difficult so see that for each partition $\mathcal{P}$ the set $\mathcal{I}_{\mathcal{P}}$ is a closed, convex, polyhedral cone which approximates $\mathcal{C}$ from the interior. Likewise, define the sets

$$\mathcal{O}_{\mathcal{P}} := \{A \in \mathcal{S} : v^T Av \geq 0 \ \text{ for all } v \in V_{\mathcal{P}}\}.$$

These sets can be shown to be closed, convex, polyhedral cones which approximate $\mathcal{C}$ from the exterior. For both inner and outer approximating cones the approximation of $\mathcal{C}$ gets monotonically better if the partitions get finer. In the limit (i.e., if the diameter $\delta(\mathcal{P}) := \max_{\{u,v\} \in E_{\mathcal{P}}} \|u - v\|$ of the partitions goes to zero), the cones $\mathcal{I}_{\mathcal{P}}$ converge to $\mathcal{C}$ from the interior, and the $\mathcal{O}_{\mathcal{P}}$ converge to $\mathcal{C}$ from the exterior.

Note that due to their polyhedrality optimizing over $\mathcal{I}_{\mathcal{P}}$ or $\mathcal{O}_{\mathcal{P}}$ amounts to solving an LP. Now replacing the cone $\mathcal{C}$ in (1) by $\mathcal{I}_{\mathcal{P}}$ and $\mathcal{O}_{\mathcal{P}}$, respectively, results in two sequences of LPs whose solutions are upper, resp. lower, bounds of the optimal value of (1). Under standard assumptions, this algorithm is provably convergent.

The performance of this method relies on suitable strategies to derive simplicial partitions $\mathcal{P}$ of the standard simplex, and in this sense the approach resembles a Branch-and-Bound algorithm. The partitioning strategy can be guided adaptively through the objective function, yielding a good approximation of $\mathcal{C}$ in those parts of the cone that are relevant for the optimization and only a coarse approximation in those parts that are not.

A drawback is that the number of constraints in the auxiliary LPs grows very quickly and the constraint systems contain a lot of redundancy. This necessitates rather involved strategies to keep the size of the systems reasonable, but nonetheless computer memory (not cpu-time) remains the limiting factor for this algorithm.

The algorithm is not adequate for general models derived from Burer's result [19], and provides only poor results for box-constrained quadratic problems. However, the method turns out to be very successful for the standard quadratic problem: while a standard global optimization solver like BARON [55] solves StQPs in 30 variables in about 1000 seconds, this method solves problems in 2000 variables in 30 seconds (on average). This shows that the copositive approach to StQPs outperforms all other available methods.

A variant of this approach can be found in [56].

Conclusion and outlook

Copositive programming is a new versatile research direction in conic optimization. It is a powerful modelling tool and allows to formulate many combinatorial as well as nonconvex quadratic problems. In the copositive formulation, all intractable constraints (binary as well as quadratic constraints) get packed entirely in the cone constraint. Studying the structure of the copositive and completely positive cones thus provides new insight to both combinatorial and quadratic problems. Though formally very similar to semidefinite programs, copositive programs are NP-hard. Nonetheless, the copositive formulations have lead to new and tighter bounds for some combinatorial problems. Algorithmic approaches to directly solve copositive and completely positive problems have been proposed and given encouraging numerical results.

Copositive optimization continues to be a highly active research field. Future research will deal with both modeling issues and algorithmic improvements. For example, it would be intersting to extend Burer's result to problems with general quadratic constraints. The now available algorithms are not successful for all copositive models, so we need other, better models for some problem classes. It will also be very interesting to see new copositivity driven cutting planes for various combinatorial problems which will emerge from a better understanding of the facial geometry of $\mathcal{C}$.

On the algorithmic side, the methods need to be improved and adapted to different problem classes. Since now a very good algorithm for StQPs is available, a natural next step is to tailor this algorithm to QPs with arbitrary linear constraints or box constraints.

References

1. Andersson L.E., Chang G.Z., Elfving T., Criteria for copositive matrices using simplices and barycentric coordinates. *Linear Algebra and its Applications* 220(1995): 9–30.
2. Anstreicher K.M., Burer S., D.C. versus copositive bounds for standard QP. *Journal of Global Optimization* 33(2005): 199–312.
3. Baston V.J., Extreme copositive quadratic forms. *Acta Arithmetica* 15(1969): 319–327.
4. Baumert L.D., Extreme copositive quadratic forms. *Pacific Journal of Mathematics* 19(1966): 197–204.
5. Baumert L.D., Extreme copositive quadratic forms II. *Pacific Journal of Mathematics* 20(1967): 1–20.
6. Berman A., Shaked-Monderer N., *Completely positive matrices*, World Scientific, 2003.
7. Berman A., Rothblum, U., A note on the computation of the cp-rank. *Linear Algebra and its Applications* 419(2006), 1–7.
8. Berman A., Hershkowitz D., Combinatorial results on completely positive matrices. *Linear Algebra and its Applications* 95(1987), 111–125.
9. Bomze I.M., Building a completely positive factorization. *Technical Report* TR-ISDS 2009-06, Department of Statistics and Decision Support Systems, University of Vienna, Austria. Online at `http://www.optimization-online.org/DB_HTML/2009/08/2381.html`
10. Bomze I.M., Linear-time copositivity detection for tridiagonal matrices and extension to block-tridiagonality, *SIAM Journal on Matrix Analysis and Applications* 21(2000): 840–848.
11. Bomze I.M., Dür M., de Klerk E., Roos C., Quist A.J., Terlaky T., On copositive programming and standard quadratic optimization problems, *Journal of Global Optimization* 18(2000): 301–320.
12. Bomze I.M., de Klerk E., Solving standard quadratic optimization problems via linear, semidefinite and copositive programming. *Journal of Global Optimization* 24(2002): 163–185.
13. Bomze I.M., Jarre F., A note on Burer's copositive representation of mixed-binary QPs. *Technical Report* TR-ISDS 2009-04, Department of Statistics and

Decision Support Systems, University of Vienna, Austria. Online at
`http://www.optimization-online.org/DB_HTML/2009/08/2368.html`

14. Bomze I.M., Jarre F., Rendl F., Quadratic factorization heuristics for copositive programming. *Technical Report* TR-ISDS 2009-08, Department of Statistics and Decision Support Systems, University of Vienna, Austria. Online at `http://www.optimization-online.org/DB_HTML/2009/10/2426.html`
15. Bomze I.M., Schachinger W., Multi-standard quadratic optimization: interior point methods and cone programming reformulation. *Computational Optimization and Applications* 45(2009): 237–256.
16. Bundfuss S., Copositive matrices, copositive programming, and applications. *Ph.D. Dissertation*, TU Darmstadt 2009. Online at `http://www3.mathematik.tu-darmstadt.de/index.php?id=483`
17. Bundfuss S., Dür M.: An adaptive linear approximation algorithm for copositive programs. *SIAM Journal on Optimization* 20(2009): 30–53.
18. Bundfuss S., Dür M.: Algorithmic copositivity detection by simplicial partition. *Linear Algebra and its Applications* 428(2008): 1511–1523.
19. Burer S., On the copositive representation of binary and continuous nonconvex quadratic programs. *Mathematical Programming* 120(2009): 479–495.
20. Burer S., Anstreicher K.M., Dür M., The difference between 5×5 doubly nonnegative and completely positive matrices. *Linear Algebra and its Applications* 431(2009): 1539–1552.
21. Cottle R.W., Habetler G.J., Lemke C.E., On classes of copositive matrices. *Linear Algebra and its Applications* 3(1970): 295–310.
22. Danninger G., A recursive algorithm for determining (strict) copositivity of a symmetric matrix. *Methods of Operations Research* 62(1990): 45–52.
23. de Klerk E., Pasechnik D.V., Approximation of the stability number of a graph via copositive programming, *SIAM Journal on Optimiaztion* 12(2002): 875–892.
24. Dickinson P.J.C., An improved characterisation of the interior of the completely positive cone. *Preprint* (2010).
25. Drew J.H., Johnson C.R., Loewy R., Completely positive matrices associated with M-matrices *Linear Algebra and Multilinear Algebra* 37(1994): 303–310.
26. Dukanovic I., Rendl F., Copositive programming motivated bounds on the stability and the chromatic numbers. *Mathematical Programming* 121(2010): 249–268.
27. Dür M., Still G., Interior points of the completely positive cone. *Electronic Journal of Linear Algebra* 17(2008): 48–53.
28. Eichfelder G., Jahn J., Set-semidefinite optimization. *Journal of Convex Analysis* 15(2008): 767–801.
29. Feng Y.Y., Li P., Criteria for copositive matrices of order four. *Linear Algebra and its Applications* 194(1993), 109–124.
30. Gvozdenović N., Laurent M., The operator Ψ for the chromatic number of a graph. *SIAM Journal on Optimization* 19(2008), 572–591.
31. Hall Jr. M., Newman M., Copositive and completely positive quadratic forms. *Proceedings of the Cambridge Philosophical Society* 59(1963): 329–33.
32. Haynsworth E., Hoffman A.J., Two remarks on copositive matrices. *Linear Algebra and its Applications* 2(1969): 387–392.
33. Hiriart-Urruty J.B., Seeger A., A variational approach to copositive matrices. Forthcoming in *SIAM Review*.
34. Hoffman A.J., Pereira F., On copositive matrices with $-1, 0, 1$ entries. *Journal of Combinatorial Theory* 14(1973): 302–309.

35. Ikramov K.D., Linear-time algorithm for verifying the copositivity of an acyclic matrix. *Computational Mathematics and Mathematical Physics* 42(2002): 1701–1703.
36. Ikramov K.D., Savel'eva N., Conditionally definite matrices. *Journal of Mathematical Sciences* 99(2000): 1–50.
37. Jarre F., Schmallowsky K., On the computation of $\mathcal{C}^*$ certificates. *Journal of Global Optimization* 45(2009): 281–296.
38. Johnson C.R., Reams R., Spectral theory of copositive matrices. *Linear Algebra and its applications* 395(2005): 275–281.
39. Kaykobad M., On nonnegative factorization of matrices. *Linear Algebra and its applications* 96(1987), 27–33.
40. Kogan N., Berman A., Characterization of completely positive graphs. *Discrete Mathematics* 114(1993): 297–304.
41. Maxfield J.E., Minc H., On the matrix equation $X'X = A$, *Proceedings of the Edinburgh Mathematical Society*, 13:125–129, 1962/1963.
42. National Bureau of Standards, Report 1818. *Quarterly Report, April through June 1952.*
43. Motzkin, T.S., Straus, E.G., Maxima for graphs and a new proof of a theorem of Turan, *Canadian Journal of Mathematics* 17(1965), 533–540.
44. Murty K.G., Kabadi S.N., Some NP-complete problems in quadratic and nonlinear programming. *Mathematical Programming* 39(1987): 117–129.
45. Natarajan K., Teo C.P., Zheng Z., Mixed zero-one linear programs under objective uncertainty: a completely positive representation. *Preprint.* Online at `http://www.optimization-online.org/DB_HTML/2009/08/2365.html`
46. Nesterov Y., Nemirovskii A., *Interior-point polynomial algorithms in convex programming*, SIAM Studies in Applied Mathematics, 13.
47. Parrilo P., Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization, *Ph.D. Dissertation*, California Institute of Technology, 2000. Available at: `http://etd.caltech.edu/etd/available/etd-05062004-055516/`
48. Peña J., Vera J., Zuluaga L., Computing the stability number of a graph via linear and semidefinite programming, *SIAM Journal on Optimization* 18(2007): 87–105.
49. Pólya G., Über positive Darstellung von Polynomen. *Vierteljahresschrift der naturforschenden Gesellschaft in Zürich* 73(1928): 141–145.
50. Povh J., Rendl F., Copositive and semidefinite relaxations of the Quadratic Assignment Problem. *Discrete Optimization* 6(2009): 231–241.
51. Povh J., Rendl F., A copositive programming approach to graph partitioning. *SIAM Journal on Optimization* 18(2007): 223–241.
52. Preisig J.C., Copositivity and the minimization of quadratic functions with nonnegativity and quadratic equality constraints. *SIAM Journal on Control and Optimization* 34(1996): 1135–1150.
53. Quist A.J., de Klerk E., Roos C., Terlaky T., Copositive relaxation for general quadratic programming. *Optimization Methods and Software* 9(1998): 185–208.
54. Sponsel J., Bundfuss S., Dür, M., Testing copositivity using semidefinitness. *Manuscript in preparation* (2009).
55. Tawarmalani M., Sahinidis N.V., Global optimization of mixed-integer nonlinear programs: A theoretical and computational study. *Mathematical Programming* 99(2004): 563–591.

56. Yıldırım E.A., On the accuracy of uniform polyhedral approximations of the copositive cone. *Preprint*. Online at http://www.optimization-online.org/DB_HTML/2009/07/2342.html